Speech planning: key models and results, part II
17 Oct. 2017

1 Clarifying terminology a bit: Keating & Shattuck-Hufnagel 2002

- phonological encoding: "generating the representation which will support an articulatory plan for an utterance, on the basis of the morphosyntactic specification of its underlying sentence and other kinds of information"
- broad sense (Levelt 1989): the whole process of retrieving & processing the word form
- narrow sense (LRM99): **"**process by which an incomplete phonological representation is completed"

3 steps of most interest to us

- **morphological** encoding
  - "retrieving the required morpheme(s)"
  - like <escort> and <-ing> if you want the present-progressive of the lemma *escort*.

- **phonological** encoding
  - activate metrical frame (σ σ́)
  - activate segments (ɛ s k ɔ ɹ t ɪ ŋ)
  - attach segments to syllable nodes in metrical frame

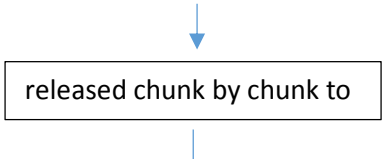
- **phonetic** encoding
  - retrieve stored gestural score for each syllable
  - (or build new one when necessary)

2 More about the scope of pre-phonological access: Wheeldon 2013 lit review

See there for references.

2.1 The need for a buffer

- **message**: non-linear; concepts assigned to thematic roles



- **process of *grammatical encoding*:** select lexical items, generate syntax, determine word order
- but sometimes a chunk is released before grammatical encoding is ready for it
  - ROSA AND JAMES – LAST NIGHT – DANCED TOGETHER
  - LAST NIGHT must be held in temporary storage

- other stuff that can require waiting
  - agreement that requires you to wait for the head
    - Spanish *l-a nuev-a generación*
      the-FEM new-FEM generation

  - idioms/conventional collocations
    - *fall/sink into disuse/oblivion*
    - you need to know the object to know which verb is the conventional one

  - binomials (not mentioned in article)
    - *bread and butter*

2.2 Evidence for clause as a unit of (maybe conceptual?) planning (or lexical access)

- planning pauses tend to be at clause boundaries
- word-swapping errors tend to be from within the same clause
- agreement errors tend to be caused by a local noun that's in the same clause
  - common: *The slogan on the posters **are** really effective*
  - uncommon: *The boy [that likes the snakes] **are** really happy*

2.3 Evidence for phrase or smaller as a unit of (maybe grammatical?) planning

- semantically related distractor word interfers with verb only in VSO clause, not SVO or SOV (German experiment)
- gaze shift to second picture happens just before picture's name is articulated
  - "we plan each item to the level of phonological encoding before moving the yes to the next object to be named, with planning progressing only slightly ahead of articulation"
  - but there may be an initial scan of all the objects that happens first—that might be for conceptual access
    - so again, the conceptual level might have a bigger planning unit

2.4 Evidence for phrase from Wheeldon & colleagues studies of speech onset latency

i.e., how long does it take to start talking?

- previous finding
  - takes longer to start saying: *[The circle and the square] move up*
  - takes less time to start saying: *[The circle] moves up and the square moves up*

  ➔ length of **subject** is crucial (doesn't matter how complex the rest of the sentence is)

- experiment: describe pictures
  - *[The dog and the hat] move above [the fork]*: longer latency
  - *[The dog] moves above [the hat and the fork]*: shorter latency

  ➔ again, looks like the **subject phrase** is the first planning unit
  - But is the effect at the conceptual, grammatical, or phonological level?
    - or even, "first two content words"
      - where *moves* doesn't count because it's always *moves* in this experiment

- replicate in Japanese
  - *[X & Y] [Z above] be*: longer latency
  - *[X] [Y & Z above] be*: shorter latency
    - that rules out the "first two content words" idea

- subject, or head of subject phrase? or actor?
  - *[the fork and the dog] are blue*: longer latency
  - *[the fork above the dog] is blue*: shorter latency

  ➔ looks like **head** is the unit
  - Japanese
    - *[[crab-above] [kettle & desk & panda] red-are*: shortest latency
    - *[crab & kettle above] [desk & panda] red-are*
    - *[crab & kettle & desk above] panda red-is*: longest latency

    ➔ you don't have to get all the way to the head of the subject before you start talking (but you do try to prepare the whole PP)
- "These findings therefore suggest that the locus of the effect is in the process of *grammatical* encoding"

2.5 Scope of lexical access

Leveltians

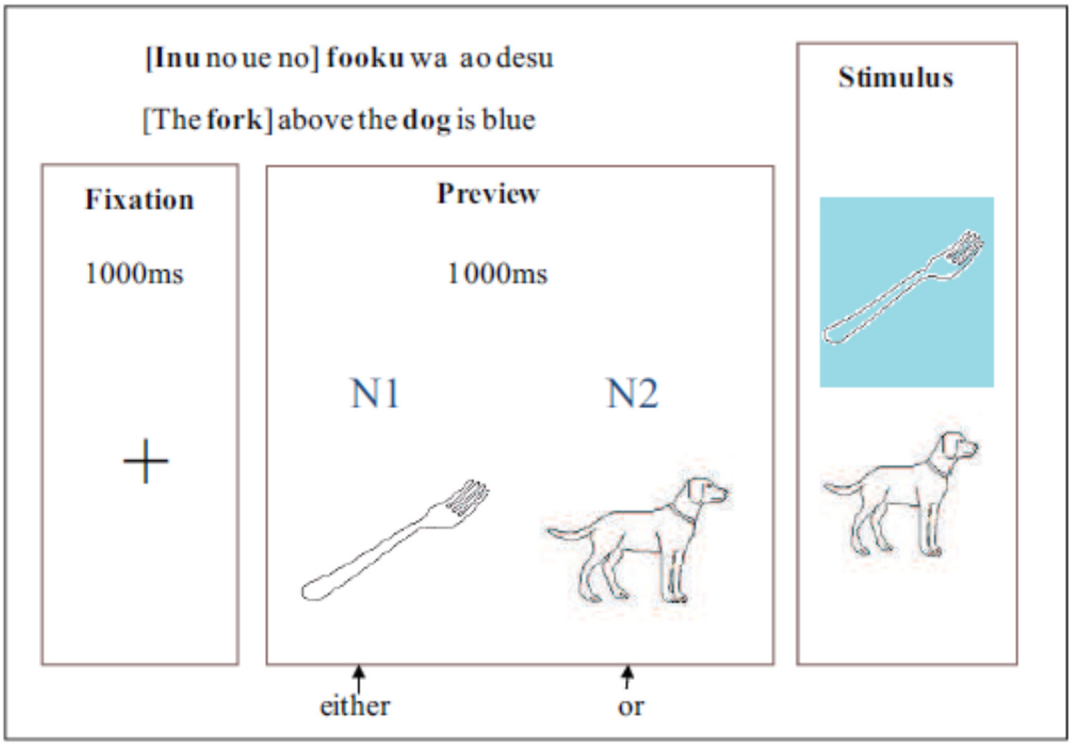
- Activate salient argument first, like actor; then you can assign it to subject
  - *The **dog** bites …*
- Or if theme is more salient, you'll end up with a passive construction.
  - *The **man** is bitten …*
- But that's not sufficient, because "grammatical convention" (i.e., the grammar!) determines whether you say "crab-above kettle" (Japanese) or "kettle above the crab" (English)
  - So the grammar has to be able to say which order to activate (or at least, say!) lexical items

Changians (e.g., Chang 2002)

- Generate syntactic structure from arguments, without needing to do lexical access first.
  - [AGENT] BITE [PATIENT]
  - Lexical selection is independent of figuring out syntax/word order
  - "the word for the concept *Anne* will be activated in the same way regardless of the thematic role it plays in the message"

Picture preview task

- "Picture preview" task: before seeing the picture you're supposed to describe (the fork above the dog is blue), you see either a blank screen, the fork, or the dog.



(p. 106)

  - seeing the first noun helps
  - seeing the second noun doesn't help
  - true for English: [the **fork** above the dog] is blue
  - also true for Japanese: [**dog**-above fork is blue]

  ➔ it's not the head (underlined) that matters, but the first noun (bold)

  ➔ supports idea that thematic structure isn't important to lexical access

- follow-up: [X&Y] are blue
  - seeing X helps, but so does seeing Y.

  ➔ "**first phrase**" is still important

2.6 Wheeldon conclusions

⇨ "more processing time is devoted to the beginning of an utterance prior to speech onset, than to the utterance as a whole" (p. 110)

⇨ phrase-sized unit, at level of grammatical encoding

⇨ "syntactic linearization processes precede lexical access processes" ← problematic for phonological effects on word order! (that's going to be my highlight)

3 More about the scope of planning in phonological processing (Damian & Dumay 2007 literature review)

- Anywhere from a segment (minimalists) to a phonological word (non-minimalists)

3.1 Studies that manipulate properties of later parts of utterance to see if they affect response latency

- Wheeldon & Levelt 1995: bisyllabic words. Frequency of second syllable matters (but see Levelt, Roelofs & Meyer 1999, p. 32, say D&D)
  → second syllable is assembled before utterance initiated
- Meyer, Roelofs & Levelt 2003: length of word matters
  → metrical frame retrieved

3.2 Studies that try to interfere with later parts of the utterance (name picture while seeing or hearing distractor word)

- Meyer & Schriefers 1991: phonological overlap is facilitatory, even at the end of the word
  → whole word is phonologically encoded before articulation begins

Some studies of phrases find small window

- Jescheniak & Schriefers 2001: German N or Det+N
  - More facilitation for just N than for Det+N (if distractor is form-related)

  → more likely to get phonologically encoded if first word

- Schriefers & Teruel 1999: German Adj+Noun
  - no phonological facilitation for noun
  - facilitation for Adj only if overlap with first syllable (overlap with second syllable weak, unreliable effect)
  - Does that mean some people start talking when they've only got a syllable ready?
    - Those who prepare only one syllable should be more likely to restart, hesitate, pause.
    - Speakers with fewer disfluent errors within adjective did show priming effect for second syllable; those with more such errors didn't
  - D&D ask: are these stable individual differences, or adaptation to the task (on that particular occasion, some participants choose to minimize error, and others to maximize speed)

Others find bigger window

- Miozzo & Caramazza 1999: Italian N or Det+N
  - Similar amount of facilitation in both cases
- Costa & Caramazza 2002: English D+Adj+N (*the green plane*)
  - facilitation of noun

- Schnur, Costa & Caramazza 2006: *The orange girl walks*
  - facilitated by distractor that form-overlaps with verb

  → **third** word already encoded, on at least some trials
- Jescheniak, Schriefers & Hantsch 2003: form overlap with noun facilitates bare N, Det+N, but not Det+Adj+Noun (instead, inhibition)

  → but that still means the Noun was at least undergoing phonological access of some kind, right?

3.3 Studies that look at duration of (parts of) utterance

- Studies that look at duration of (parts of) utterance
  - "response durations of relatively short phrases are not affected by manipulations such as relatedness in PWI [picture/word interference], providing additional evidence for a relatively large extent of advance planning prior to articulation"
  - cf. Kello, Plaut & MacWhinney: Stroop task (say what color a word is printed in, but it's a color word! I can demonstrate this on the board) shows longer latencies when word & color mismatch, but not longer utterance duration. When response deadline was added, now you get Stroop effect on durations too.
    - interpretation: normally you don't start talking till after planning is done (so planning difficult affects latency, but duration is then unaffected), but if you're forced to respond faster, you can start talking before you're ready, but duration will be stretched out because you're still planning while talking
    - Damian 2003 failed to replicate this

4 More about adding strain on working memory (Slevc 2011)

4.1 Working memory in speech planning

- As Wheeldon discussed, you can end up with situations where lemma has been retrieved but isn't ready to be produced, and has to be held in working memory

- "Release valve" theory of working memory
  - capacity is limited! don't keep stuff in working memory longer than necessary
  - if *book* is ready, say *The pirate gave the **book** to the monk.*
  - if *monk* is ready, say *The pirate gave the **monk** the book.*
    - predicts that more strain on working memory → more pressure to offload stuff as soon as possible

- Retrieval-manager theory of working memory
  - there's a big set of activated stuff, and working memory's job is to manage what is getting retrieved
  - Whatever's easier for it to retrieve will get used first (*book* or *monk*)

- What promotes saying a word earlier in the sentence? Previous findings:
  - imageability, animacy, semantic priming
  - givenness in the discourse
  - but not phonological content → this research is probably not about phonological encoding

4.2 Experiment: describe pictures that involve datives

- Experimenter asks "What's going on with the pirate?"
  - There are various pictures on the wall
  - Participant (only—experimenter can't see screen) then sees BOOK or MONK on the screen
    - This should increase its accessibility
  - Both these cues (*pirate* and *book/monk*) are needed for participant to know which picture to describe

- interference
  - half of trials also show thm two unrelated words that participant has to remember and be able to say back after the trial is over
  - release-valve theory's prediction
    - pressure to use accessible stuff sooner should get **stronger** under strain
  - retrieval-controlling theory's prediction
    - working memory should be worse at retrieving stuff when under strain, **weakening** tendency to produce the accessible word sooner

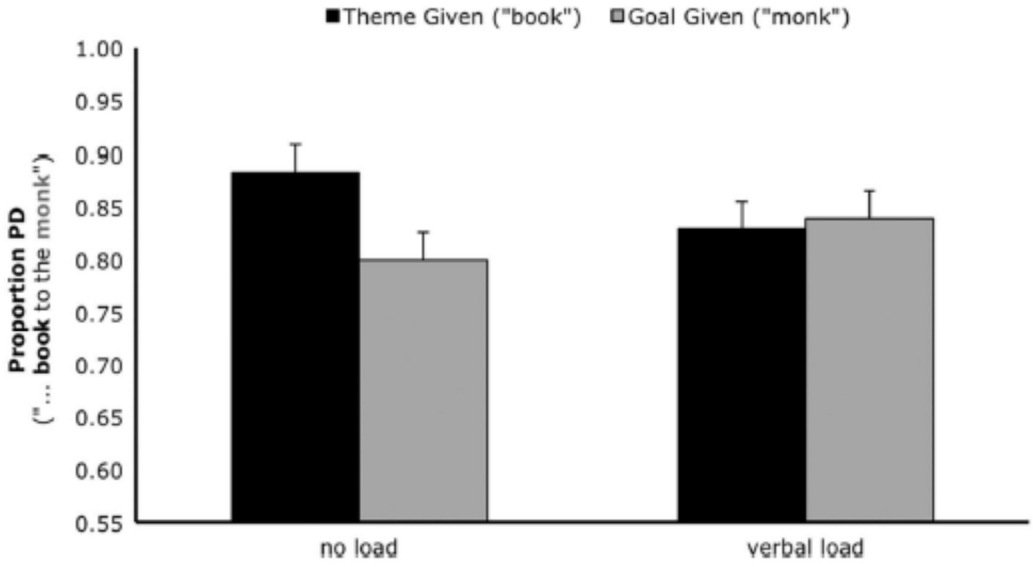


Figure 1. Proportion of prepositional dative (PD) sentences produced in Experiment 1 as a function of cue type (i.e., which postverbal argument was made accessible: theme or goal) and WM load (none or two-word load). Note that the *y*-axis does not start at zero. Error bars represent standard error.

(p. 1506)

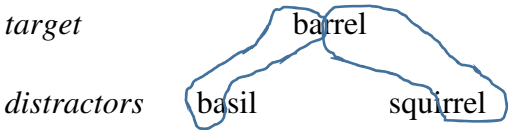
⇨ effect (use the cued word earlier) goes away under verbal working-memory load

- Exp. 2: rather than cue, info is **given or new**
  - *What's going on with the pirate and the book?*
  - *What's going on with the pirate and the monk?*
  - effect (use given word earlier) is weaker under verbal WM load

- Exp. 3: **verbal vs. non-verbal load**
  - sometimes the « load » is remembering a pattern of dots, rather than words
  - effect is strong under spatial load, weak under verbal load
  - is the spatial load just harder?
    - No: in letter-rotation task, spatial load causes worse performance than verbal load

5 More about effects of time pressure on <u>phonological planning</u>: Damian & Dumay 2007

- Name a picture
  - simultaneously, a distractor word is shown.
  - distractor overlaps either beginning sounds or end sounds of target (or is unrelated)



  - half of participants get a warning message and tone if they're taking too long on a trial
- Exp 1. just a noun *(barrel)*
  - faster if distractor word begins or ends with same sounds as target picture name

  → establishes that overlapping sounds are helpful (shorter response latency)

- Exp. 2. Det+Adj+Noun *(the green clock)*
  - still faster if distractor ends with same sounds as target noun
  - even though noun is second content word

  → Noun is at least form-retrieved, even if not phonologically encoded, before speaking begins

- Exp. 3. No distractor word, Det+Adj+Noun.
  - Faster if Adj and Noun start with same sounds
    - *the blue bottle*
    - (ending overlap not tested)
  - Even more time pressure
    - shorter deadline
    - instructions changed to make errors sound more tolerable (if you're not making at least some mistakes, you're not responding fast enough). Again, means (at least beginning of) Noun is being phonologically planned before utterance initiation. (Trend for less facilitation under deadline condition, but apparently not significant)

- All three: time pressure ("response deadline") makes response faster, but doesn't weaken any of the priming effects.
  - "pressure to provide a swift response does no reduce the **scope** of phonological planning"
  - cf. Wagner, Jescheniak & Schriefers 2010
    - slow responders planned more before starting to talk
    - but that was a post-hoc analysis of who was fast and who was slow
    - speed wasn't manipulated

⇨ "no reason to believe that speakers are ever able to initiate a response before they have phonologically encoded at least a single word. Instead, our findings suggest that advance planning encompasses at least two phonological words, even under time pressure"

- They speculate that natural speech could be different, because it's harder (you have to so much more, at the conceptual and syntactic levels)

6 Our highlights of Keating & Shattuck-Hufnagel 2002, Wheeldon 2013

mine for Wheeldon (see above for K & S-H): There are more segment-switching errors across adjacent PWords than within a PWd . Why is that?

Let's assign one whiteboard to each of our two readings, and write our highlights on the board

7 Reading plan for speech errors & OCP

1. (Berg & Abd-El-Jawad 1996)	speech errors are different in Arabic than in English & German, because non-concatenative morphology → different process of phonological encoding	_____ will present
2. (Frisch 2004)	producing or perceiving repetition is hard → motivation for OCP	_____ will present
3. (Hansson 2001)	consonant harmony as phonologized speech errors	_____ will present
4. (Rose & Walker 2004)	focus on section 3.1, conclusion; long-distance consonant agreement as "grammaticized avoidance of combinations that may present production difficulties"	_____ will present
5. (McAllister Byun & Inkelas 2014)	performance pressures in child speech can become phonologized (for an individual child)	_____ will present
6. (Rose & King 2007)	experiments to elicit speech errors in languages with strong OCP (Amharic, Chaha)	_____ will present
7. (Walker, Hacopian & Taki 2002)	inducing speech errors to compare to long-distance nasal harmony	_____ will present
8. (Walker 2007)	larger-scale speech-error study on nasal harmony	_____ will present

I haven't thought of a good point/counterpoint format for these papers, but I do have an idea for our directionality papers

References

Berg, Thomas & Hassan Abd-El-Jawad. 1996. The unfolding of suprasegmental representations: a cross-linguistic perspective. *Journal of Linguistics* 32(2). 291–324. doi:10.1017/S0022226700015905.

Chang, Franklin. 2002. Symbolically speaking: a connectionist model of sentence production. *Cognitive Science* 26(5). 609–651.

Costa, Albert & Alfonso Caramazza. 2002. The Production of Noun Phrases in English and Spanish: Implications for the Scope of Phonological Encoding in Speech Production. *Journal of Memory and Language* 46(1). 178–198. doi:10.1006/jmla.2001.2804.

Damian, Markus F. & Nicolas Dumay. 2007. Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language* 57(2). 195–209. doi:10.1016/j.jml.2006.11.001.

Frisch, Stefan. 2004. Language processing and segmental OCP effects. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*. Cambridge University Press.

Hansson, Gunnar Olafur. 2001. The phonologization of production constraints: evidence from consonant harmony. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *CLS 37: The main session. Papers from the 37th meeting of the Chicago Linguistic Society*, 187–200. Chicago: Chicago Linguistic Society.

Jescheniak, Jörg D. & Herbert Schriefers. 2001. Priming effects from phonologically related distractors in picture-word interference. *The Quarterly Journal of Experimental Psychology Section A* 54(2). 371–382. doi:10.1080/713755981.

Jescheniak, Jörg D., Herbert Schriefers & Ansgar Hantsch. 2003. Utterance format affects phonological priming in the picture-word task: implications for models of phonological encoding in speech production. *Journal of Experimental Psychology. Human Perception and Performance* 29(2). 441–454.

Keating, Patricia & Stefanie Shattuck-Hufnagel. 2002. A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics* 101. 112–156.

Levelt, Willem J. M., Ardi Roelofs & Antje S. Meyer. 1999. A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences* 22(01). 1–38.

McAllister Byun, Tara & Sharon Inkelas. 2014. Child consonant harmony and phonologization of performance errors. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.), *NELS 43: Proceedings of the Forty-Third Annual Meeting of the North East Linguistic Society, Volume I*. Amherst, MA: GLSA.

Meyer, A. S. & H. Schriefers. 1991. Phonological facilitation in picture-word interference experiments : effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of experimental psychology. Learning, memory, and cognition* 17(6). 1146–1160.

Meyer, Antje S, Ardi Roelofs & Willem J. M Levelt. 2003. Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language* 48(1). 131–147. doi:10.1016/S0749-596X(02)00509-0.

Miozzo, M. & A. Caramazza. 1999. The selection of determiners in noun phrase production. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 25(4). 907–922.

Rose, Sharon & Lisa King. 2007. Speech Error Elicitation and Co-occurrence Restrictions in two Ethiopian Semitic Languages. *Language and Speech* 50(4). 451–504. doi:10.1177/00238309070500040101.

Rose, Sharon & Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80(3). 475–532.

Schnur, Tatiana T., Albert Costa & Alfonso Caramazza. 2006. Planning at the Phonological Level during Sentence Production. *Journal of Psycholinguistic Research* 35(2). 189–213. doi:10.1007/s10936-005-9011-6.

Schriefers, H. & E Teruel. 1999. Phonological Facilitation in the Production of Two-word Utterances. *European Journal of Cognitive Psychology* 11(1). 17–50. doi:10.1080/713752301.

Slevc, L. Robert. 2011. Saying what's on your mind: Working memory effects on sentence production. *Journal of experimental psychology. Learning, memory, and cognition* 37(6). 1503–1514. doi:10.1037/a0024350.

Wagner, Valentin, Jörg D. Jescheniak & Herbert Schriefers. 2010. On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 36(2). 423–440. doi:10.1037/a0018619.

Walker, Rachel. 2007. Nasal and oral consonant similarity in speech errors: Exploring parallels with nasal consonant harmony. *Language and Cognitive Processes* 22(7). 1073–1113. doi:10.1080/01690960701317859.

Walker, Rachel, Narineh Hacopian & Mariko Taki. 2002. Nasal consonant speech errors: implications for "similarity" and nasal harmony at a distance. *First Pan-American/Iberian Meeting on Acoustics*. Cancun.

Wheeldon, L. R. & W. J. M. Levelt. 1995. Monitoring the Time Course of Phonological Encoding. *Journal of Memory and Language* 34(3). 311–334. doi:10.1006/jmla.1995.1014.

Wheeldon, Linda. 2013. Producing spoken sentences: the scope of incremental planning. In S Fuchs, M Weirich, D Pape & P Perrier (eds.), *Speech production and perception volume 1: speech planning and dynamics*, 93–114. Peter Lang.